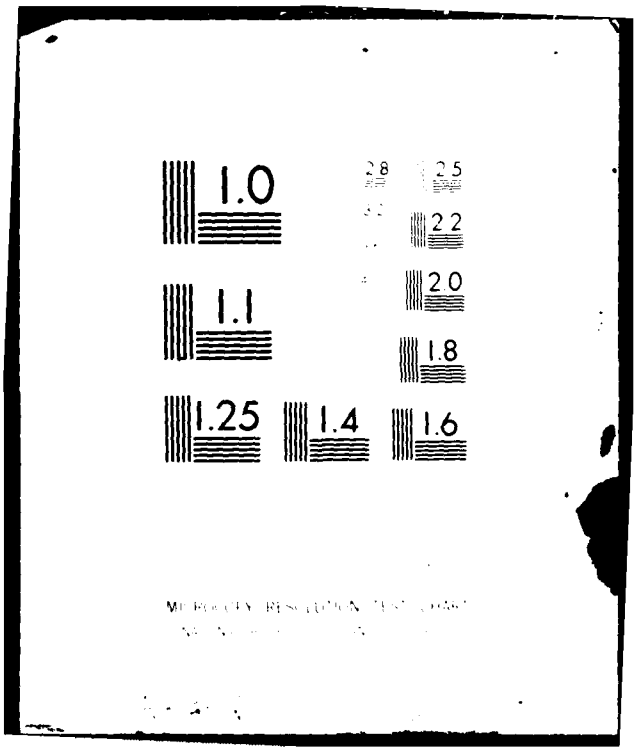
AD-A110 298 YALE UNIV NEW HAVEN CONN DEPT OF STATISTICS F/G 12/1
FINAL REPORT ON CONTRACT F49620-79-C-0164.(U)
NOV 81 D POLLARD F49620-79-C-0164
UNCLASSIFIED AFOSR-TR-81-0878 NL


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5

November 30, 1981

FINAL REPORT

AFSC contract F49620-79-C-0164

Principal Investigator: David Pollard

The main achievement of the research conducted under this contract was the sharpening of techniques for studying the asymptotic behaviour of the empirical measure P_n constructed from independent sampling on a distribution P. The best results obtained used symmetrization and combinatorial methods to bound the deviations of P_n from P. This gave central limit theorems for $n^{\frac{1}{2}}(P_n - P)$, which generalized the classical Donsker theorem for empirical distribution functions [2], [4]. These were applied to determine the limiting distribution of some statistics expressible as functionals $T_n P_n$ of the empirical measure, as in the central limit theorem for the optimal k-means cluster centres [5]. It turned out that these methods also had applications to the communications problem of quantization. The invited paper [6] explores one such connection.

Gary Oehlert and Adrian Baddeley were supported as research assistants during parts of the contract period. Gary completed a Ph.D. on the asymptotic behaviour of alternative estimators of the mean; he now holds the position of Assistant Professor of Statistics at Princeton. Adrian contributed a lot of geometric know-how to the k-means problem. Some of our theoretal expectations of the rotational instability of these estimators are now being detected for practical quantization algorithms.

As a by-product of this work and related graduate courses, a new approach to the study of weak convergence methods has emerged. Pursuit of this approach and a widening of the applications of the empirical process tools -- to cover rank tests, estimation from censored data, and other problems in quantization -- seems the natural next step.

Publications:

[1] Strong consistency of k-means clustering. Ann. Statistics 9 (1981), 135-140.

[2] Limit theorems for empirical processes. Z. Wahrscheinlichkeitstheorie und verw. Geb. 47 (1981), 181-195.

[3] Beyond the heuristic approach to Kolmogorov-Smirnov theorems. In: Festschrift fur P.A.P. Moran (1982) -- to appear.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER: AFOSR-TR- 81-0878 / 2. GOVT ACCESSION NO.: AD-A110298	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): FINAL REPORT on Contract F49620-79-C-0164	5. TYPE OF REPORT & PERIOD COVERED: Final - 7/1/79 to 9/30/81
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): David Pollard	8. CONTRACT OR GRANT NUMBER(s): F49620-79-C-0164
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Yale University, Dept of Statistics, New Haven, CT 06520	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: F 51102 ; 2304/A6
11. CONTROLLING OFFICE NAME AND ADDRESS: Mathematical & Information Sciences Directorate, Air Force Office of Scientific Research, Bolling AFB DC 20332	12. REPORT DATE: Nov. 30, 1981
	13. NUMBER OF PAGES: Two (2)
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release;
distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Empirical processes, central limit theorem, quantization, k-means clustering.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A central limit theorem for empirical processes was proved. Applications to the electrical engineering problem of quantization and the statistical k-means clustering procedure were explored.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

10 2 80

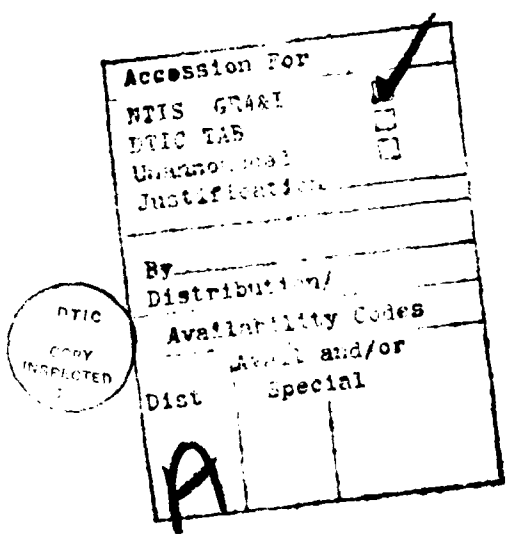
[4] A central limit theorem for empirical processes. J. Australian Math. Soc. (1982) -- to appear.

[5] A central limit theorem for k-means cluster centres. Ann. Probability (198?) -- to appear.

[6] Quantization and the method of k-means. IEEE Transactions on Information Theory (Special issue on quantization) 28 (1982) -- to appear.

PROPERTY INVENTORY

No property items were purchased under this contract, so there is no property to inventory.




FILMED
2-8